



**QUEEN'S
UNIVERSITY
BELFAST**

On the performance of zero-forcing processing in multi-way massive MIMO relay networks

Duc Ho, C., Ngo, H. Q., Matthaiou, M., & Duong, T. Q. (2017). On the performance of zero-forcing processing in multi-way massive MIMO relay networks. *IEEE Communications Letters*, 21(4), 849-852.
<https://doi.org/10.1109/LCOMM.2017.2648795>

Published in:
IEEE Communications Letters

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights
Copyright 2017 IEEE.
This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

On the Performance of Zero-Forcing Processing in Multi-Way Massive MIMO Relay Networks

Chung Duc Ho, Hien Quoc Ngo, Michail Matthaiou, and Trung Q. Duong

Abstract—We consider a multi-way massive multiple-input multiple-output relay network with zero-forcing processing at the relay. By taking into account the time-division duplex protocol with channel estimation, we derive an analytical approximation of the spectral efficiency. This approximation is very tight and simple which enables us to analyze the system performance, as well as, to compare the spectral efficiency with zero-forcing and maximum-ratio processing. Our results show that by using a very large number of relay antennas and with zero-forcing technique, we can simultaneously serve many active users in the same time-frequency resource, each with high spectral efficiency.

I. INTRODUCTION

Multi-way relay networks are relevant for many applications, such as data transfer in multimedia teleconference and data exchange between sensor nodes and data fusion centers in wireless communications [1]. Due to the multiplexing gain, the spectral efficiency of multi-way relay networks is much larger than that of two-way or one-way relay networks. Therefore, during the past years, multi-way relay networks have attracted considerable research interest [2]. On a parallel avenue, massive multiple-input multiple-output (MIMO) has also attracted a significant amount of research interest from both academia and industry [3]. In massive MIMO, hundreds of antennas are deployed at the base station to serve simultaneously tens of users. With simple linear processing techniques, such as maximum-ratio (MR) or zero-forcing (ZF) processing, massive MIMO can offer huge spectral and energy efficiency [4]. Thus, massive MIMO combined with multi-way relaying technique is a strong candidate for the next-generation wireless communication systems.

Recently, there have been some works in multi-way massive MIMO relay systems [5], [6]. These systems can offer all benefits of both massive MIMO and multi-way relaying technologies, and hence, are expected to offer very high spectral efficiency. In particular, in [5], the authors show that by using very large antenna arrays at the relay together with ZF processing, the system performance can improve significantly. Furthermore, [6] shows that the transmit power of each user and/or the relay can be made inversely proportional to the number of relay antennas, while maintaining a required quality of service. However, these works assume perfect channel state information (CSI) at the relay and users. In practice, especially in massive MIMO systems, the impact of channel estimation should be taken into consideration. In [7], the authors analyze

the performance of multi-way massive MIMO systems with imperfect CSI and MR processing at the relay. To the best of the authors' knowledge, there is no work on ZF processing with imperfect CSI in literature, partially due to the difficulty in manipulating products of Wishart matrices.

In this paper, we investigate a multi-way massive MIMO relay network with ZF processing and imperfect CSI. The relay estimates the channels via uplink pilots and the minimum mean square error (MMSE) scheme. Then, it uses these channel estimates and the ZF technique to combine and beamform the signals to all users. We derive an approximate closed-form expression for the spectral efficiency. This approximation is very tight and enables us to further analyze the performance of the considered system.

Notation : The superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ stand for the transpose, conjugate, and Hermitian, respectively. The notations $\mathbb{E}\{\cdot\}$ and $\text{Var}\{\cdot\}$ are the expectation and the variance operators, respectively. Furthermore, $[\mathbf{A}]_k$ or \mathbf{a}_k denotes the k -th column of matrix \mathbf{A} .

II. MULTI-WAY MASSIVE MIMO RELAY MODEL

We consider a multi-way relaying massive MIMO network which includes one relay station and K users.¹ The relay station is equipped with M antennas, and each user has a single antenna ($M > K$). In this system, each user wants to communicate with $K - 1$ other users with the aid of the relay. We assume that the direct links (user-to-user links) are absent due to large path loss and/or severe shadowing.

Let $\mathbf{G} \in \mathbb{C}^{M \times K}$ be the channel matrix from the K users to the relay, which includes the small-scale fading and the large-scale fading and is modelled as

$$\mathbf{G} = \mathbf{H}\mathbf{D}^{1/2}, \quad (1)$$

where $\mathbf{H} \sim \mathcal{CN}(0, \mathbf{I}_M)$ represents the small-scale fading, and $\mathbf{D} \in \mathbb{C}^{K \times K}$ is a diagonal matrix containing the large-scale fading coefficients whose k -th diagonal element is denoted by β_k .

The transmission protocol is the same as the one in [7]. More precisely, the data exchange between all the K users is done via time-division duplex (TDD) operation. With TDD operation, each coherence interval is divided into three phases: channel estimation, multiple-access, and broadcast phases.

A. Channel Estimation Phase

All the K users simultaneously send pilot sequences to the relay. The relay then estimates the channels to all users through receiving pilots. Let T and τ be the lengths of each coherence interval and the training duration (in symbols), respectively, with $T > \tau$. We assume that the pilots used by the K users

Manuscript received November 26, 2016. The associate editor coordinating the review of this paper and approving it for publication was C. Masouros.

The authors are with the School of Electronics, Electrical Engineering and Computer Science, Queens University Belfast, Belfast BT3 9DT, U.K. (e-mail: {choduc01, m.matthaiou, trung.q.duong}@qub.ac.uk, hien.ngo@liu.se). H. Q. Ngo is also with the Department of Electrical Engineering (ISY), Linköping University, 581 83 Linköping, Sweden.

This work was supported by project no. 3811/QD-UBND, Binh Duong government, Vietnam. The work of H. Q. Ngo was supported by the Swedish Research Council (VR) and ELLIT. The work of M. Matthaiou was supported in part by the EPSRC under grant EP/P000673/1. The work of T. Q. Duong was supported by the U.K. Royal Academy of Engineering Research Fellowship under Grant RF1415\14\22.

¹It would be more practical to consider multi-cell setups. Unfortunately, if we consider multi-cell setups, the system model becomes too complicated to analyze. Note that our results can be regarded as an upper bound of what is actually achieved in multi-cell setups. If pilot reuse scheme is applied, then this upper bound is very tight [8].

are pairwise orthogonal. This requires $\tau \geq K$. We denote by P_p the normalized transmit signal-to-noise ratio (SNR) per pilot symbol. Then, the MMSE channel estimate of \mathbf{G} can be represented as [7]

$$\hat{\mathbf{G}} = \mathbf{G} - \mathbf{E}, \quad (2)$$

where \mathbf{E} is the estimation error matrix, which is independent of $\hat{\mathbf{G}}$. Furthermore, $\hat{\mathbf{G}} \sim \mathcal{CN}(0, \hat{\mathbf{D}})$ and $\mathbf{E} \sim \mathcal{CN}(0, \mathbf{D}_E)$, where $\hat{\mathbf{D}}$ and \mathbf{D}_E are diagonal matrices whose (k, k) -th elements are $\sigma_k^2 = \frac{\tau P_p \beta_k^2}{\tau P_p \beta_{k+1} + 1}$, and $\sigma_{e,k}^2 = \beta_k - \sigma_k^2$, respectively.

B. Multiple-Access Phase

After sending the pilot sequences' phase, all the users simultaneously send their data to the relay. Let x_k , where $\mathbb{E}\{|x_k|^2\} = 1$, is the signal transmitted from the k -th user. Then, the relay sees

$$\mathbf{y}_R = \sqrt{P_u} \mathbf{G} \mathbf{x} + \mathbf{n}_r, \quad (3)$$

where P_u is the normalized transmit SNR, $\mathbf{x} \triangleq [x_1, \dots, x_K]^T$, and $\mathbf{n}_r \sim \mathcal{CN}(0, \mathbf{I}_M)$ is the AWGN vector at the relay. Then, the relay uses the channel estimate and ZF technique to combine the received signals from all M antennas as

$$\tilde{\mathbf{y}}_R = \mathbf{W}^T \mathbf{y}_R, \quad (4)$$

where \mathbf{W}^T is the ZF receiver given by [4]

$$\mathbf{W}^T = \left(\hat{\mathbf{G}}^H \hat{\mathbf{G}} \right)^{-1} \hat{\mathbf{G}}^H. \quad (5)$$

C. Broadcast Phase

To send all signals to K users, the relay spends $K-1$ time-slots. In the t -th time-slot, the relay aims to send x_{k+t} to user k (if $k+t > K$, then x_{k+t} is set to be x_{k+t-K}). Thus, the transmit signal vector at the relay for the t -th time-slot is

$$\mathbf{s}_R^{(t)} = \sqrt{\alpha^{(t)}} \mathbf{A} \mathbf{\Pi}^{(t)} \tilde{\mathbf{y}}_R, \quad t = 1, 2, \dots, K, \quad (6)$$

where $\mathbf{\Pi}^{(t)} \in \mathbb{C}^{K \times K}$ is the permutation matrix at the t -th time-slot given by [5, Eq. (17)], \mathbf{A} is the ZF precoding matrix expressed as

$$\mathbf{A} = \hat{\mathbf{G}}^* \left(\hat{\mathbf{G}}^T \hat{\mathbf{G}}^* \right)^{-1}, \quad (7)$$

and $\alpha^{(t)}$ is chosen to satisfy the power constraint at the relay,

$$\mathbb{E} \left\{ \left\| \mathbf{s}_R^{(t)} \right\|^2 \right\} = P_r. \quad (8)$$

Denote $\mathbf{B}^{(t)} \triangleq \mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T \mathbf{G}$, and $\mathbf{C}^{(t)} \triangleq \mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T$. Then $\mathbf{s}_R^{(t)}$ in (6) can be rewritten as

$$\mathbf{s}_R^{(t)} = \sqrt{\alpha^{(t)}} P_u \mathbf{B}^{(t)} \mathbf{x} + \sqrt{\alpha^{(t)}} \mathbf{C}^{(t)} \mathbf{n}_r. \quad (9)$$

Plugging (9) into (8), we have

$$\alpha^{(t)} = \frac{P_r}{P_u \mathbf{Q}_1^{(t)} + P_u \mathbf{Q}_2^{(t)} + \mathbf{Q}_3^{(t)}}, \quad (10)$$

where

$$\mathbf{Q}_1^{(t)} \triangleq \mathbb{E} \left\{ \text{Tr} \left[\left(\mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T \hat{\mathbf{G}} \right) \left(\mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T \hat{\mathbf{G}} \right)^H \right] \right\}, \quad (11)$$

$$\mathbf{Q}_2^{(t)} \triangleq \mathbb{E} \left\{ \text{Tr} \left[\left(\mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T \mathbf{E} \right) \left(\mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T \mathbf{E} \right)^H \right] \right\}, \quad (12)$$

$$\mathbf{Q}_3^{(t)} \triangleq \mathbb{E} \left\{ \text{Tr} \left[\left(\mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T \right) \left(\mathbf{A} \mathbf{\Pi}^{(t)} \mathbf{W}^T \right)^H \right] \right\}. \quad (13)$$

With the transmitted signal given in (9), the K users receive

$$\begin{aligned} \mathbf{y}_u^{(t)} &= \mathbf{G}^T \mathbf{s}_R^{(t)} + \mathbf{n}_u^{(t)} \\ &= \sqrt{\alpha^{(t)}} P_u \mathbf{G}^T \mathbf{B}^{(t)} \mathbf{x} + \sqrt{\alpha^{(t)}} \mathbf{G}^T \mathbf{C}^{(t)} \mathbf{n}_r + \mathbf{n}_u^{(t)}, \end{aligned} \quad (14)$$

where $\mathbf{n}_u \sim \mathcal{CN}(0, \mathbf{I}_K)$ is the AWGN vector at the users.

III. SPECTRAL EFFICIENCY ANALYSIS

We derive a closed-form expression for the spectral efficiency of the transmission in the first time-slot. The same analysis can be done for other time-slots. Note that, hereafter, we set $k+1 = 1$ if $k = K$ and set $k-1 = K$ if $k = 1$.

By using the bounding technique in [9], the received signal at the k -th user $y_{u,k}^{(1)}$ is expressed as:

$$y_{u,k}^{(1)} = \underbrace{\sqrt{\alpha^{(1)}} P_u \mathbb{E} \left\{ \mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)} \right\}}_{\text{desired signal}} x_{k+1} + \underbrace{\tilde{N}_k^{(1)}}_{\text{effective noise}}, \quad (15)$$

where

$$\begin{aligned} \tilde{N}_k^{(1)} &\triangleq \sqrt{\alpha^{(1)}} P_u \left(\mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)} - \mathbb{E} \left\{ \mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)} \right\} \right) x_{k+1} \\ &+ \sqrt{\alpha^{(1)}} P_u \sum_{\substack{i=1 \\ i \neq (k+1)}}^K \mathbf{g}_k^T \mathbf{b}_i^{(1)} x_i + \sqrt{\alpha^{(1)}} \mathbf{g}_k^T \mathbf{C}^{(1)} \mathbf{n}_r + n_{u,k}^{(1)}. \end{aligned} \quad (16)$$

The worst-case Gaussian noise yields an achievable spectral efficiency for the k -th user, which is given as

$$\text{SE}_k^{(1)} = \left(\frac{T-\tau}{T} \right) \left(\frac{K-1}{K} \right) \log_2 \left(1 + \frac{\alpha^{(1)} P_u \left| \mathbb{E} \left\{ \mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)} \right\} \right|^2}{\text{Var} \left(\tilde{N}_k^{(1)} \right)} \right). \quad (17)$$

To derive the spectral efficiency in closed-form, we need to compute $\mathbb{E} \left\{ \mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)} \right\}$ and $\text{Var} \left(\tilde{N}_k^{(1)} \right)$. From the independence between $\hat{\mathbf{G}}$ and \mathbf{E} , we have

$$\mathbb{E} \left\{ \mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)} \right\} = \mathbb{E} \left\{ \mathbf{g}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \hat{\mathbf{g}}_{k+1} \right\} = 1. \quad (18)$$

Since $\tilde{N}_k^{(1)}$ has a complicated form which includes matrix inversions and multiplications of Wishart matrices, we cannot obtain an exact closed-form of $\text{Var} \left(\tilde{N}_k^{(1)} \right)$. However, thanks to the law of large numbers (for large M), we can obtain the following approximation.

Theorem 1: As $M \rightarrow \infty$, the spectral efficiency (17) can be approximated as

$$\text{SE}_k^{(1)} \rightarrow \left(\frac{T-\tau}{T} \right) \left(\frac{K-1}{K} \right) \log_2 \left(1 + \frac{\alpha^{(1)} P_u}{\alpha^{(1)} P_u \sum_{i=1}^K \mathcal{I}_{k,i} + \alpha^{(1)} \mathcal{J}_k + 1} \right), \quad (32)$$

where

$$\alpha^{(1)} \triangleq \frac{M(M-K)P_r}{M P_u \sum_{k=1}^K \frac{1}{\sigma_k^2} + P_u \sum_{k=1}^K \sigma_{e,k}^2 \varrho + \varrho}, \quad (33)$$

$$\mathcal{I}_{k,i} \triangleq \frac{M \sigma_{i-1}^2 \sigma_{e,i}^2 + M \sigma_{k+1}^2 \sigma_{e,k}^2 + \sigma_{k+1}^2 \sigma_{i-1}^2 \sigma_{e,k}^2 \sigma_{e,i}^2 \varrho}{M(M-K) \sigma_{i-1}^2 \sigma_{k+1}^2}, \quad (34)$$

$$\mathcal{J}_k \triangleq \frac{M + \sigma_{k+1}^2 \sigma_{e,k}^2 \varrho}{M(M-K) \sigma_{k+1}^2}, \text{ and } \varrho \triangleq \sum_{k'=1}^K \frac{1}{\sigma_{k'}^2 \sigma_{k'+1}^2}. \quad (35)$$

Proof: See Appendix VI-B. ■

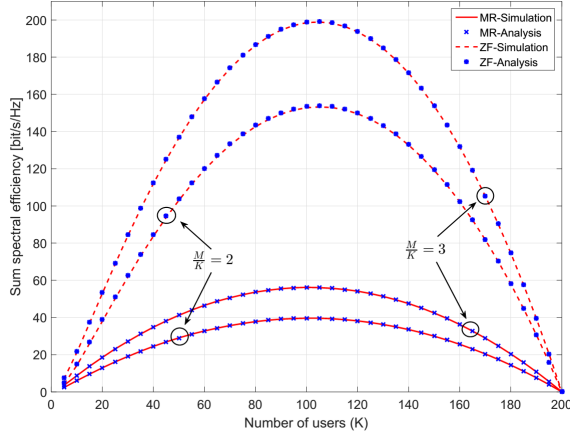


Fig. 1. Sum spectral efficiency versus the number of users K . We choose $T = 200$, $P_u = P_p = 0$ dB, $P_r = 10$ dB, and $\beta_k = 1$.

IV. NUMERICAL RESULTS

We consider the sum spectral efficiency as our performance metric. The sum spectral efficiency is defined as

$$SE_{\text{sum}} = \sum_{k=1}^K SE_k^{(1)} \quad \text{bit/s/Hz.} \quad (36)$$

For the first example, we assume that $\beta_k = 1$, and choose $T = 200$, $P_u = P_p = 0$ dB, $P_r = 10$ dB. Figure 1 compares the performance of multi-way massive MIMO systems for ZF and MR processing with different K , while the ratio M/K is kept fixed. For MR processing, we used the results in [7, Eq. (26)]. Clearly, the simulated spectral efficiency and the approximate one match perfectly. At a small K (low inter-user interference) and large K (large channel estimation overhead), the spectral efficiencies of ZF and MR processing are comparable. However, when $K = 20$ –180, ZF significantly outperforms MR processing. Interestingly, regardless of the ratio M/K , the sum spectral efficiency is maximum when K is around 100. Furthermore, when M/K increases, the inter-user interference reduces, and hence, the sum spectral efficiency increases.

We next consider a more practical scenario where users are located uniformly at random inside a disk with the diameter of 1000 m. The large-scale fading is modelled as $\beta_k = \frac{z_k}{1+(d_k)^\nu}$, where z_k is the log-normal random variable with standard deviation of 8 dB, $\nu = 4$ denotes the path-loss exponent, and d_k is the distance between the k -th user and the relay. Furthermore, the normalized transmit SNRs P_r , P_u and P_p can be calculated by dividing these powers by the noise power N_0 . In this example, we choose $N_0 = -120$ dB. We consider 2 cases: Case-1 corresponds to ($P_u = P_p = 0.2$ W, $P_r = 1$ W), and Case-2 corresponds to ($P_u = P_p = 0.1$ W, $P_r = 0.5$ W). Figure 2 shows the cumulative distribution of the sum spectral efficiency for ZF and MR processing. We can see that, at high transmit power (Case-1), the spectral efficiency of ZF processing is higher than the one of MR processing and vice versa at low transmit power. Furthermore, compared with MR processing, the spectral efficiency of ZF processing is less concentrated around its median.

V. CONCLUSION

We have investigated a multi-way massive MIMO relay network with ZF processing and imperfect CSI. We derived

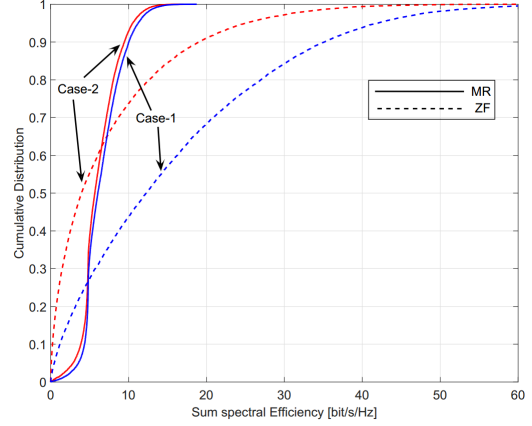


Fig. 2. Cumulative distribution of the sum spectral efficiency. Here $M = 100$, $K = 20$.

a new tractable approximate closed-form expression for the spectral efficiency. At a large number of relay antennas, the inter-user interference and noise reduces significantly, and hence, the system can offer a substantial sum spectral efficiency. Furthermore, we showed that, for most of the cases (particularly at high SNRs), ZF processing offers a higher spectral efficiency than MR processing does.

VI. APPENDICES

A. Preliminary Results

Lemma 1: Let $\mathbf{X} \in \mathbb{C}^{M \times K}$, $M > K$. Each row of \mathbf{X} is $\mathcal{CN}(0, \mathbf{D})$, where \mathbf{D} be a diagonal matrix. Furthermore, let $\hat{\mathbf{D}} \in \mathbb{C}^{K \times K}$ be another diagonal matrix. Then, we have

$$\mathbb{E} \left\{ \text{Tr} \left[\hat{\mathbf{D}} (\mathbf{X}^H \mathbf{X})^{-1} \right] \right\} = \frac{1}{M-K} \sum_{k=1}^K \frac{[\hat{\mathbf{D}}]_{kk}}{[\mathbf{D}]_{kk}}. \quad (37)$$

Proof: By expressing $\text{Tr} \left[\hat{\mathbf{D}} (\mathbf{X}^H \mathbf{X})^{-1} \right] = \sum_{k=1}^K \frac{[\hat{\mathbf{D}}]_{kk}}{[\mathbf{D}]_{kk}} [\mathcal{W}^{-1}]_{kk}$, where \mathcal{W} is a $K \times K$ central Wishart matrix of M degrees of freedom, and using [10, Lemma 2.10], we obtain (37). ■

Lemma 2: Let $\mathbf{A} \in \mathbb{C}^{M \times M}$, and $\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I}_M)$. Then,

$$\mathbb{E} \left\{ |\mathbf{x}^T \mathbf{A} \mathbf{x}|^2 \right\} = \text{Tr}(\mathbf{A} \mathbf{A}^H) + \text{Tr}(\mathbf{A} \mathbf{A}^*). \quad (38)$$

Proof: To obtain (38), we first express $\mathbf{x}^T \mathbf{A} \mathbf{x}$ as $\sum_{m=1}^M \sum_{m'=1}^M a_{mm'} x_m x_{m'}$, and then use the identities $\mathbb{E} \{ |x_m|^4 \} = 2$ and $\mathbb{E} \{ |x_m x_{m'}|^2 \} = 1$, for $m \neq m'$. ■

B. Proof of Theorem 1

1) *Derivation of $\alpha^{(1)}$:* From (10), to compute $\alpha^{(1)}$ we need to compute $\mathbf{Q}_1^{(1)}$, $\mathbf{Q}_2^{(1)}$, and $\mathbf{Q}_3^{(1)}$. The substitution of (5) and (7) into (11) yields

$$\mathbf{Q}_1^{(1)} = \mathbb{E} \left\{ \text{Tr} \left(\hat{\mathbf{G}}^T \hat{\mathbf{G}}^* \right)^{-1} \right\} = \frac{1}{M-K} \sum_{k=1}^K \frac{1}{\sigma_k^2}, \quad (39)$$

where in the last equality we have used Lemma 1.

To compute $\mathbf{Q}_2^{(1)}$, we substitute (5) and (7) into (12) to obtain

$$\mathbf{Q}_2^{(1)} = \sum_{k=1}^K \sigma_{e,k}^2 \mathbb{E} \left\{ \text{Tr} \left[\mathbf{\Pi}^{(1)} (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} (\mathbf{\Pi}^{(1)})^H (\hat{\mathbf{G}}^T \hat{\mathbf{G}}^*)^{-1} \right] \right\}.$$

From the law of large numbers, we have that $\hat{\mathbf{G}}^H \hat{\mathbf{G}} \rightarrow M \hat{\mathbf{D}}$, and hence, \mathbf{Q}_2 can be approximated as

$$\begin{aligned} \mathbf{Q}_2^{(1)} &\rightarrow \sum_{k=1}^K \sigma_{e,k}^2 \mathbb{E} \left\{ \text{Tr} \left[\mathbf{\Pi}^{(1)} (M \hat{\mathbf{D}})^{-1} (\mathbf{\Pi}^{(1)})^H (\hat{\mathbf{G}}^T \hat{\mathbf{G}}^*)^{-1} \right] \right\} \\ &= \frac{\varrho}{M(M-K)} \sum_{k=1}^K \sigma_{e,k}^2, \end{aligned} \quad (40)$$

where again we have used Lemma 1 to obtain the last equality. Similarly, we obtain

$$\mathbf{Q}_3^{(1)} \rightarrow \frac{\varrho}{M(M-K)}. \quad (41)$$

Substituting (39), (40), and (41) into (10), we obtain (33).

2) *Derivation of $\mathbb{V}\text{ar}(\tilde{N}_k^{(1)})$* : From (16), we have

$$\begin{aligned} \mathbb{V}\text{ar}(\tilde{N}_k^{(1)}) &= \alpha^{(1)} P_u \mathbb{V}\text{ar}(\mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)}) + \alpha^{(1)} P_u \mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_k^{(1)} \right|^2 \right\} \\ &+ \alpha^{(1)} P_u \sum_{\substack{i=1 \\ i \neq (k,k+1)}}^K \mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_i^{(1)} \right|^2 \right\} + \alpha^{(1)} \mathbb{E} \left\{ \left\| \mathbf{g}_k^T \mathbf{C}^{(1)} \right\|^2 \right\} + 1. \end{aligned} \quad (42)$$

a) Compute $\mathbb{V}\text{ar}(\mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)})$: By expressing the true channel as the sum of the channel estimate plus the channel estimation error, we obtain

$$\mathbb{V}\text{ar}(\mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)}) = \mathbf{V}_1 + \mathbf{V}_2 + \mathbf{V}_3, \quad (43)$$

where

$$\mathbf{V}_1 \triangleq \mathbb{E} \left\{ \left| \hat{\mathbf{g}}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \mathbf{e}_{k+1} \right|^2 \right\}, \quad (44)$$

$$\mathbf{V}_2 \triangleq \mathbb{E} \left\{ \left| \mathbf{e}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \hat{\mathbf{g}}_{k+1} \right|^2 \right\}, \quad (45)$$

$$\mathbf{V}_3 \triangleq \mathbb{E} \left\{ \left| \mathbf{e}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \mathbf{e}_{k+1} \right|^2 \right\}. \quad (46)$$

The term \mathbf{V}_1 can be computed as

$$\begin{aligned} \mathbf{V}_1 &= \sigma_{e,k+1}^2 \mathbb{E} \left\{ \left\| \hat{\mathbf{g}}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \right\|^2 \right\} \\ &= \sigma_{e,k+1}^2 \mathbb{E} \left\{ \left[\left(\hat{\mathbf{G}}^H \hat{\mathbf{G}} \right)^{-1} \right]_{k+1,k+1} \right\} = \frac{\sigma_{e,k+1}^2}{(M-K)\sigma_{k+1}^2}. \end{aligned}$$

Similarly, we obtain $\mathbf{V}_2 = \frac{\sigma_{e,k}^2}{(M-K)\sigma_k^2}$, and $\mathbf{V}_3 = \frac{\sigma_{e,k}^2 \sigma_{e,k+1}^2}{M(M-K)} \varrho$. Therefore,

$$\mathbb{V}\text{ar}(\mathbf{g}_k^T \mathbf{b}_{k+1}^{(1)}) = \mathcal{I}_{k,k+1}. \quad (47)$$

b) Compute $\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_k^{(1)} \right|^2 \right\}$: By expressing $\mathbf{g}_k = \hat{\mathbf{g}}_k + \mathbf{e}_k$, and using the fact that $\hat{\mathbf{g}}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \hat{\mathbf{g}}_k = 0$, we get

$$\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_k^{(1)} \right|^2 \right\} = \mathbb{E} \left\{ |\mathbf{I}_1 + \mathbf{I}_2 + \mathbf{I}_3|^2 \right\}, \quad (48)$$

where $\mathbf{I}_1 \triangleq \hat{\mathbf{g}}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \mathbf{e}_k$, $\mathbf{I}_2 \triangleq \mathbf{e}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \hat{\mathbf{g}}_k$, and $\mathbf{I}_3 \triangleq \mathbf{e}_k^T \mathbf{A} \mathbf{\Pi}^{(1)} \mathbf{W}^T \mathbf{e}_k$. Since \mathbf{I}_1 , \mathbf{I}_2 , and \mathbf{I}_3 are mutually uncorrelated, we obtain

$$\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_k^{(1)} \right|^2 \right\} = \mathbb{E} \left\{ |\mathbf{I}_1|^2 \right\} + \mathbb{E} \left\{ |\mathbf{I}_2|^2 \right\} + \mathbb{E} \left\{ |\mathbf{I}_3|^2 \right\}. \quad (49)$$

Similarly to the derivation of \mathbf{V}_1 , we have

$$\mathbb{E} \left\{ |\mathbf{I}_1|^2 \right\} = \frac{\beta_k - \sigma_k^2}{(M-K)\sigma_{k+1}^2}, \quad \mathbb{E} \left\{ |\mathbf{I}_2|^2 \right\} = \frac{\beta_k - \sigma_k^2}{(M-K)\sigma_{k-1}^2}. \quad (50)$$

Next, we compute $\mathbb{E} \left\{ |\mathbf{I}_3|^2 \right\}$. Using (38) from Lemma 2, and the law of large numbers, we obtain

$$\begin{aligned} \mathbb{E} \left\{ |\mathbf{I}_3|^2 \right\} &= \sigma_{e,k}^4 \mathbb{E} \left\{ \text{Tr} \left[\mathbf{\Pi}^{(1)} (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} (\mathbf{\Pi}^{(1)})^H (\hat{\mathbf{G}}^T \hat{\mathbf{G}}^*)^{-1} \right] \right\} \\ &+ \sigma_{e,k}^4 \mathbb{E} \left\{ \text{Tr} \left[\mathbf{\Pi}^{(1)} (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} (\mathbf{\Pi}^{(1)})^* (\hat{\mathbf{G}}^T \hat{\mathbf{G}}^*)^{-1} \right] \right\} \\ &\rightarrow \frac{\sigma_{e,k}^4 \varrho}{M(M-K)}. \end{aligned} \quad (51)$$

Substituting (50), and (51) into (49), we get

$$\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_k^{(1)} \right|^2 \right\} \rightarrow \mathcal{I}_{k,k}. \quad (52)$$

c) Compute $\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_i^{(1)} \right|^2 \right\}$, where $i \neq k, k+1$: Following a similar methodology as in the derivation of $\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_k^{(1)} \right|^2 \right\}$, we obtain

$$\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_i^{(1)} \right|^2 \right\} \rightarrow \mathcal{I}_{k,i}. \quad (53)$$

d) Compute $\mathbb{E} \left\{ \left\| \mathbf{g}_k^T \mathbf{C}^{(1)} \right\|^2 \right\}$: By replacing \mathbf{g}_k with $\hat{\mathbf{g}}_k + \mathbf{e}_k$ together with Lemma 1 and the law of large numbers, as in the derivation of $\mathbb{E} \left\{ \left| \mathbf{g}_k^T \mathbf{b}_k^{(1)} \right|^2 \right\}$, we obtain

$$\mathbb{E} \left\{ \left\| \mathbf{g}_k^T \mathbf{C}^{(1)} \right\|^2 \right\} = \frac{M + \sigma_{k+1}^2 \sigma_{e,k}^2 \varrho}{M(M-K)\sigma_{k+1}^2}. \quad (54)$$

Substituting (47), (52), (53), and (54) into (42) yields (32).

REFERENCES

- [1] D. Gündüz, A. Yener, A. Goldsmith, and H. V. Poor, "The multiway relay channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 51–63, Jan. 2013.
- [2] A. Amah and A. Klein, "Non-regenerative multi-way relaying with linear beamforming," in *Proc. IEEE PIMRC*, Sep. 2009, pp. 1843–1847.
- [3] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [5] G. Amarasingh, C. Tellambura, and M. Ardakani, "Multi-way MIMO amplify-and-forward relay networks with zero-forcing transmission," *IEEE Trans. Wireless Commun.*, vol. 61, no. 12, pp. 4847–4863, Dec. 2013.
- [6] G. Amarasingh and H. V. Poor, "Multi-way amplify-and-forward relay networks with massive MIMO," in *Proc. IEEE PIMRC*, Sep. 2014, pp. 595–600.
- [7] C. D. Ho, H. Q. Ngo, M. Matthaiou, and T. Q. Duong, "Multi-way massive MIMO with maximum-ratio processing and imperfect CSI," [Online]. Available: <https://arxiv.org/pdf/1611.01042.pdf>.
- [8] T. Marzetta, E. Larsson, H. Yang, and H. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [9] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multi-pair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.
- [10] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.